

VIVES

2012

**DISCUSSION PAPER**

**34**

# Estimating Lock Congestion

*Jo Reynaerts*

# Estimating Lock Congestion

Jo Reynaerts<sup>\*,a,b</sup>

<sup>a</sup>*Steunpunt Ondernemen & Regionale Economie (STORE)*

<sup>b</sup>*VIVES and CES, Faculty of Economics & Business, KU Leuven*

This version: 30 August 2012

## Abstract

We consistently estimate the parameters of 600 and 1200 ft lock congestion functions by exploiting the geographical variation in demand patterns for individual locks along the Upper Mississippi River system of locks and dams. In addition to avoiding endogeneity bias, we control for lock-specific heterogeneity in estimating a fixed-effects regression model that relates lock congestion to lock usage and lock characteristics. Using a panel data set spanning the years 1993–2010, we find significant empirical evidence of the existence of a quadratic lock congestion function for 600 ft technology in the sense of BPR (1964). While 1200 ft technology on the other hand seemingly operates under free-flow conditions, the presence of an auxiliary chamber significantly mitigates congestion, with the effect being more pronounced for higher levels of traffic. Unscheduled lock outages and non-commercial traffic attenuate lock congestion.

**Key words:** lock congestion, simultaneity bias, consistency, fixed-effects estimation

**JEL codes:** R4, R41

---

\*Address of correspondence: Naamsestraat 61/3550, B-3000 Leuven, Belgium. E-mail: [Jo.Reynaerts@kuleuven.be](mailto:Jo.Reynaerts@kuleuven.be). I thank Koen Deconinck, Stefaan Decramer, Klaus Desmet, Joep Konings, Damiaan Persyn, Patrick Van Cayseele, Eddy Van de Voorde and Thijs Vandemoortele for their comments and suggestions. This paper was presented at the World Conference on Transport Research Society – Special Interest Group 2 Maritime Transport & Ports (WCTRS-SIG 2) *Key Developments in the Port and Maritime Sector*, University of Antwerp, 21–22 May 2012. Note: to present dynamic graphics this paper is best displayed in Adobe Reader.

# 1 Introduction

With its 29 locks located along 1350 kilometers of navigable river length, the Upper Mississippi river system of locks and dams forms the backbone of the transportation of various goods and commodities in the U.S. As an environmental friendly and low-cost means of transportation, the Upper Mississippi is of vital importance to the agricultural markets of the American Midwest. Unfortunately, most of the locks in the system are no longer up to the task of dealing with the tonnages and flows characteristic of modern-day commercial navigation; at the time of writing, most of the locks are 70 years and older.<sup>1</sup> Consequently, maintenance costs run high and outages are frequent, causing delay and congestion throughout the system.

The last decade has witnessed considerable debate between proponents and adversaries of infrastructural improvements to the current system on whether and to what extent such enhancements would be economically desirable. As a response, the U.S. Army Corps of Engineers (USACE), charged with maintenance and operation of the locks and dams, issued a research program designed to tackle and answer of that question, the conclusions of which can be consulted online, see [USACE \(2012\)](#).<sup>2</sup> Despite the efforts to systematically assess the welfare effects of the proposed infrastructural works, surprisingly little (if any) work has concentrated on the relationship between barge traffic and congestion costs, or the effect of lock characteristics on congestion. A precise mapping of this relationship can help answer the question for example which of the currently available technologies (600 or 1200 ft) to implement in the event a lock (or a series of locks) is up for replacement.

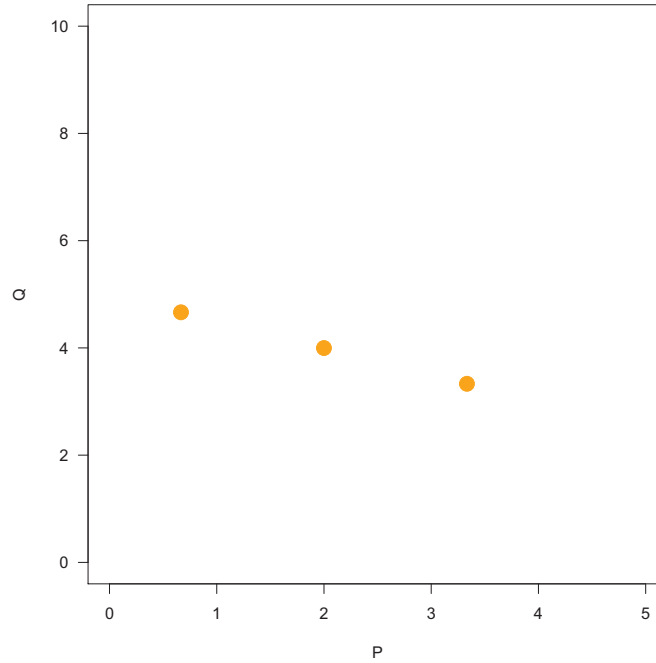
In this paper we statistically estimate the relationship between lockages (the number of barges locked) and congestion (the average delay time), taking into account both lock characteristics and lock heterogeneity. The fundamental question that arises however is whether one can consistently estimate lock congestion on the basis of observed data. Ever since the seminal work of [Working \(1925\)](#) and [Working \(1927\)](#) it is known that in the absence of demand and supply shifters, the use of observed prices and quantities might result in *simultaneity bias*, producing biased (and inconsistent) estimates of the parameters of both the demand and supply function ([Leamer, 1981](#)), see [Figure 1](#) for an illustration. In this respect, the fundamental question translates to whether one can consistently estimate the parameters  $\theta \in \Theta$  of the lock congestion function (LCF)  $f(\mathbf{x}; \theta)$  given a sequence of observations  $\{\text{barges}_{\ell ct}, \text{costs}_{\ell ct}\}_{t=1}^T$  on lock costs and the number of barges locked.<sup>3</sup> The answer to this question is not without importance: an exact measure of the impact of traffic demand on congestion, duly accounting for

---

<sup>1</sup>With the exception of Upper and Lower St. Anthony Falls and Lock 27 (1963, 1959 and 1953 respectively), only a single new lock has been constructed, i.e. the Melvin Price lock (1990, 1994).

<sup>2</sup>See in particular [USACE \(2002\)](#), [Boland et al. \(2004\)](#) and [USACE \(2004\)](#).

<sup>3</sup>In keeping with econometric tradition, we define consistency as  $\mathbb{E}[\hat{\theta}] = \theta_0$ , or as  $\lim_{T \rightarrow \infty} \hat{\theta}_t = \theta_0$ , where  $\mathbb{E}[\cdot]$  is the expectations operator, and  $\theta_0$  is the true parameter vector. In common language, consistency means that the estimated parameters are unbiased, or that they converge to their true value as the sample size grows.



**Figure 1:** Demand and supply simultaneity. In the absence of data on demand and supply shifters, observed market prices and quantities  $\{P_t, Q_t\}_{t=1}^T$  yield insufficient information to consistently estimate the parameters of either demand or supply. In the example generated above, market data suggests a downward-sloping demand curve, whereas these data are equilibrium values generated by the intersection of demand  $D_t$  and supply  $S_t$  at  $t = 1, \dots, 3$ .

lock characteristics, might serve as a direct input into many of the feasibility studies on lock capacity expansion that started this very same research. Additionally, estimation of demand for inland waterway transportation, or they might serve as parameters for link performance functions in traffic assignment problems (TAP) with competing modes of transportation, see e.g. [Sheffi \(1985, Ch. 9\)](#).

To tackle the simultaneity problem, we consistently estimate the parameters of 600- and 1200 foot lock-technology congestion functions by exploiting the geographical differences in demand patterns for individual locks along the Upper Mississippi River system of locks and dams. In analogy with the econometric literature on demand and supply estimation ([Hayashi, 2000, Ch. 3](#)), we argue that this spatial diversity in lock usage (demand) traces out the (fixed) technology-specific lock congestion function (supply) to produce unbiased estimates of the parameters of the lock congestion function. Applying the methodology, we control for lock-specific heterogeneity in estimating a fixed-effects regression model that relates lock congestion to lock usage and lock characteristics, using a panel data set spanning the years 1993–2010. We find significant empirical evidence of the existence of a quadratic lock congestion function for 600 ft technology in the sense of BPR (1964). While 1200 ft technology on the other hand seemingly operates under free-flow conditions, the presence of an auxiliary chamber significantly mitigates congestion, with the effect being more pronounced for higher levels of traffic. Unscheduled lock outages

and non-commercial traffic attenuate lock congestion.

The remainder of this paper is structured as follows: in Section 2 we provide a brief survey of related literature in economics and transportation research on the estimation of congestion (or link performance) functions, and point out relevant differences where necessary. Section 3 describes the Upper Mississippi River system of locks with respect to lock characteristics and lock usage. The numbers and figures presented in this section serve to illustrate (i) our main argument of using geographical differences in demand patterns for individual locks as an identification strategy, and (ii) the choice to separately estimate the parameters of the LCF for each of the main locking technologies (600 vs. 1200 ft). The specification of the regression function and corresponding estimation results are detailed in Section 4. Section 5 concludes.

## 2 Literature Review

In the field of economics, academic interest in estimating congestion originates in the seminal works by Walters (1961), Johnson (1964) and Dewees (1979). Interest in these papers is centered mainly on computing and estimating the optimal level of taxes and tolling for the purpose of reconciling private and social marginal costs of highway travel in a bid to relieve traffic congestion. In this paper we analyze lock congestion and abstract from the congesting-relieving effects of tolling and/or taxes.

A specific focus on estimating the travel-time relationship in the field of transportation is to be found in Taylor (1977), Rose and Raymond (1992), Rose (1994), and more recently, Hazelton and Pueschel (1999) and Russo and Vitetta (2011). Contrary to the economic analysis of congestion effects, these authors explicitly address the question on how to estimate the parameters of link performance functions for road travel, which is also the main research question in this paper. We however additionally control for and chart the influence of other factors that might influence (lock) congestion; the cited papers only discuss the direct relation between flow and congestion.

While there exists a specific literature on inland waterway transportation and the analysis of lock congestion in particular, either the subject and methods of analysis deal with congestion as a queuing process, see for example USACE (1993), or the economic analysis of the relationship between barge traffic, lock characteristics and congestion costs is absent, as mentioned in the introduction (Boland *et al.*, 2004).

## 3 The Upper Mississippi River System of Locks

### 3.1 Lock Characteristics

The main purpose of the Upper Mississippi river system of locks and dams is to enable commercial navigation, providing navigable depth of at least 9 ft for barges shipped along the river. As such, locks are operated by U.S. Army Corps of Engineers (US-

**Table 1:** Breakdown by chamber length and width (in ft)

	54	105	108	110	Total
400	3	0	0	0	3
500	0	0	0	1	1
600	0	9	11	2	22
1200	0	0	1	2	3

**Table 2:** Breakdown by chamber length (in ft) and number of chambers

	1	2	Total
400	2	1	3
500	1	0	1
600	20	2	22
1200	1	2	3
Total	24	5	

ACE) personnel 24 hours a day, 7 days a week all year round on the basis of a rotating shift schedule. Although lock technologies along the the Upper Mississippi come in different varieties (as measured in ft), both in length {360, 400, 500, 600, 1200} and width {54, 110}, the standard configuration is  $110 \times 600$  ft ( $33.5\text{m} \times 182.9\text{m}$ ). Some of the locks dispose of a smaller auxiliary chamber; for these locks the standard configuration is either  $110 \times (600, 360)$  or  $110 \times (1200, 600)$  ft. Table 1 presents a breakdown of the number of locks by length and width while Table 2 contains a breakdown by chamber length and the number of chambers.<sup>4</sup> Summarized, the Upper Mississippi river features twenty-two 600 ft locks (two of which have an auxiliary 360 ft chamber) and three 1200 ft locks (two of which have an auxiliary 360 ft chamber). The smaller 400 and 500 ft locks mainly serve recreational vessels and handle no commercial navigation; we accordingly remove these from further analysis.<sup>5</sup>

<sup>4</sup>Note that depending on the source consulted, chamber width can further be classified according to the criterion of chamber useable width, or CHMBUW following [USACE \(2011d\)](#), see Table 1. For purposes of estimation we make no such distinction and consider all 600 and 1200 ft locks to have an equal width of 110 ft.

<sup>5</sup>For an overview of specific locks characteristics including age, gate type and location, see Table 7 in the Appendix; note that locks are consecutively ordered from the last (downstream) to the first (upstream) lock in the system.

**Table 3:** Summary statistics for lock costs (in minutes) and barges locked (in thousands)

	Costs	Barges
Observations	522	522
Minimum	4.80	0.87
Maximum	585.00	84.17
Mean	112.21	20.61
Median	91.20	17.14
Variance	7249.37	261.53
Standard deviation	85.14	16.17

### 3.2 Lock Usage

Table 3 presents summary statistics for the entire system of locks on the number of barges locked (in thousands) and corresponding lock costs, defined as the sum of delay time and processing time (in minutes).<sup>6</sup> Note that due to the aggregate nature of the data provided by [USACE \(2011a\)](#), statistics on lock costs are yearly averages. While some of the specificities of the data are already apparent from the numbers in the table, for example the wide variety in lock costs (between 4.8 and 585 minutes), most of the features are masked by virtue of their aggregate nature.

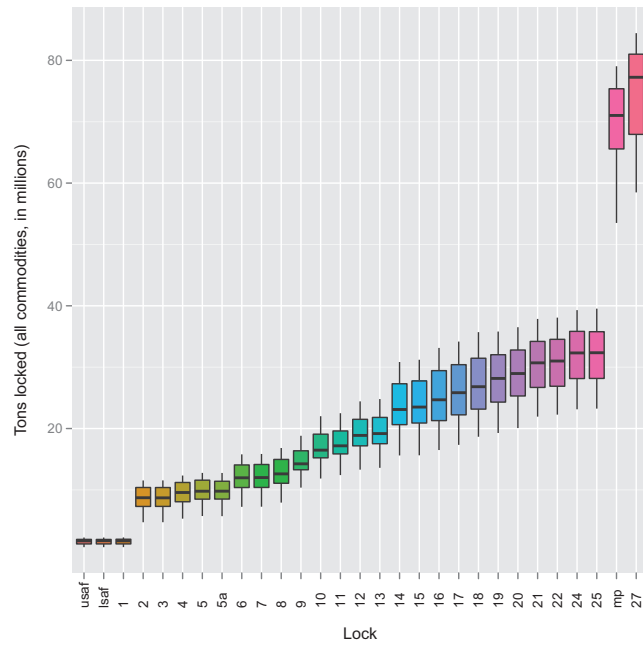
Combining lock characteristics and lock costs, Figure 2 shows the boxplot of the total number of tons locked (in millions of tons) for each lock in the system between 1993 and 2010.<sup>7</sup> Locks are ordered by their respective place in the Upper Mississippi river system of locks and dams. The boxplot clearly indicates different patterns of usage along the latitudinal dimension of the system: downstream locks handle more traffic. Figures 3 and 4 however fail to reveal a one-to-one mapping between lockages and lock costs, indicating a more intricate relationship between traffic and congestion on the Upper Mississippi river system of locks; these figures at the least indicate that length and/or the number of chambers influence (mitigate) lock congestion, or that shippers facing such conditions face less congestion (on average).<sup>8</sup>

Therefore, if we are to disentangle the factors influencing lock congestion, we need to

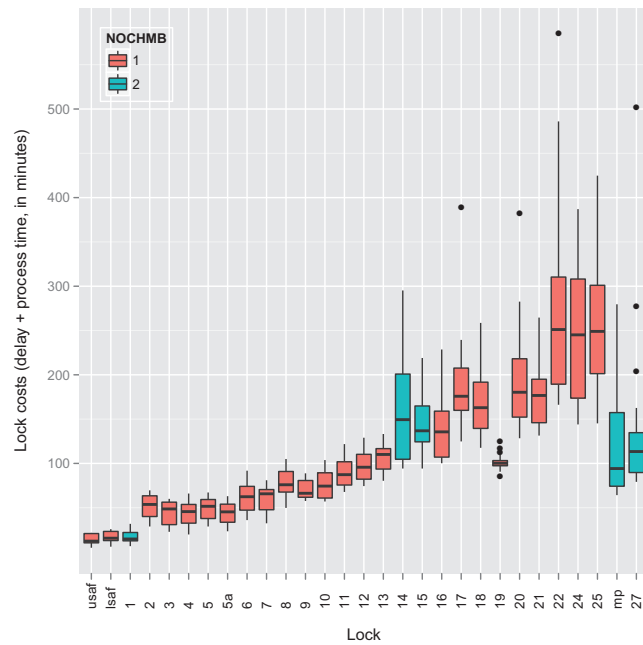
<sup>6</sup>We subscribe the economic point of view here that the relevant cost to the shipper in deciding whether or not to use the lock is the total cost incurred, the sum of delay and processing time.

<sup>7</sup>In total, lockages for nine commodity groups are monitored, see Table 6 in the Appendix. In terms of tons locked, metal products and agricultural produce constitute the largest groups. The composition of traffic handled by locks (not shown here) differs along the the latitudinal dimension of the system; downstream locks handle more of commodity groups 30 (chemicals) and 60 (agricultural produce).

<sup>8</sup>Although of different magnitude, separate boxplots for delay time and processing time (not shown here) preserve the relationship shown in Figure 2. The most important component (in magnitude) however is delay time. Figures 8, 9 and 10 in the appendix further indicate that (1) most of the variance in lock congestion (delay + process time) occurs in single-chamber 600 ft locks, (2) this variance is mainly due to delay at these locks, and (3) process costs are significantly lower for double-chamber 1200 ft locks.

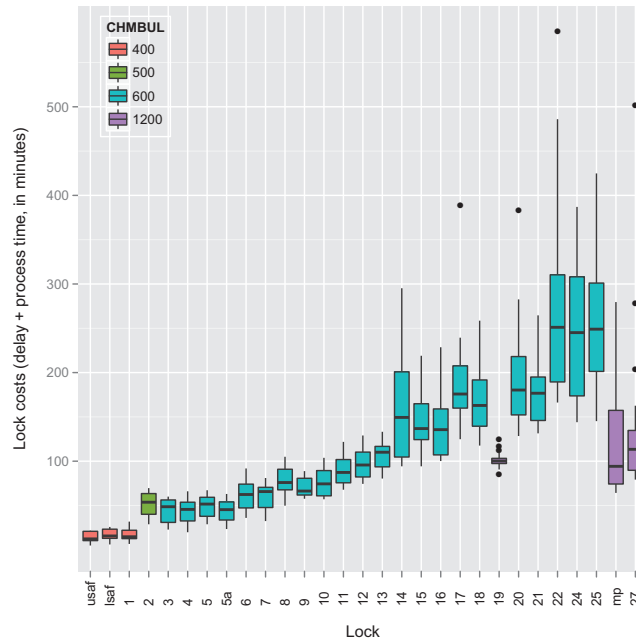


**Figure 2:** Tons locked by lock 1993–2010, all commodities (in millions of tons).



**Figure 3:** Lock costs by lock and number of chambers, 1993–2010. Single-chamber locks are presented in red, double-chamber locks in blue.





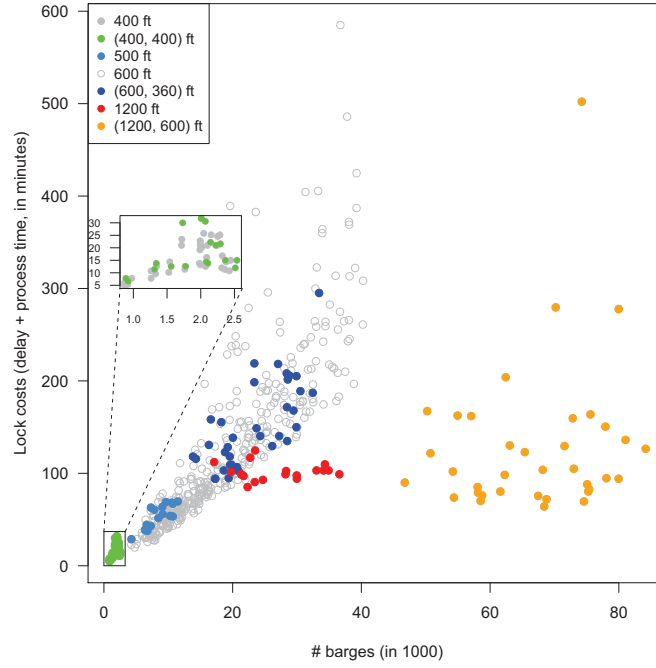
**Figure 4:** Lock costs by lock and of chamber useable length (CHMBUL), 1993–2010. The box-plot is indicative of less congestion for 1200 ft locking technology.

account for the presence of auxiliary chambers and the difference between 600 and 1200 ft lock technology. As descriptive data alone might not be conclusive, or hide other factors not accounted for, we turn to regression analysis for further analysis. Figure 5, a scatterplot of lock costs against the number of barges locked for all locks in the system, offers more evidence in favor of a regression approach and provides a number of additional observations that we will include in the specification of the regression equation (1):

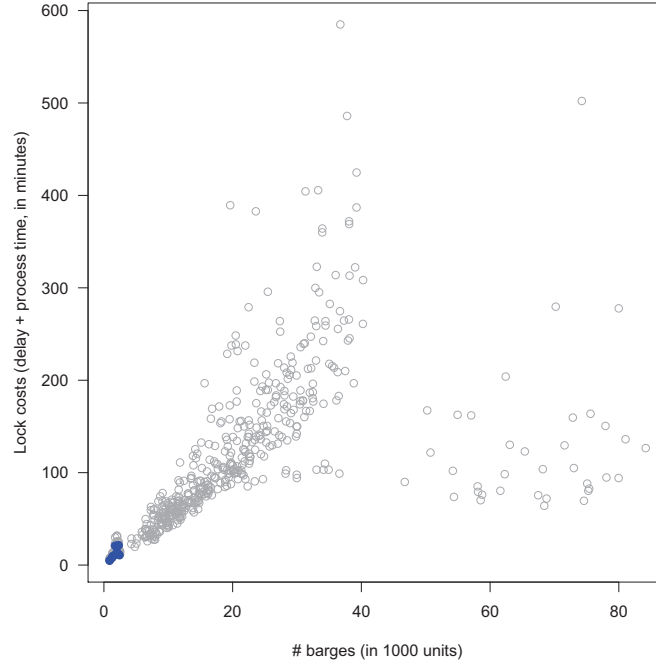
- (i) the scatterplot for the standard 600 ft configuration clearly suggests a quadratic relationship between flow and congestion;
- (ii) the presence of a second chamber does not offset this relationship for 600 ft technology, and
- (iii) at face value (and barring the influence of outliers), there appears to be no systematic relationship between barges and costs for the 1200 ft technology.

Additionally, and crucial to the argument raised in the introduction to mitigate simultaneity bias, Figure 5 also highlights the geographical diversity in demand patterns, at least for 600 ft technology, in that downstream locks face more traffic and experience more congestion, implying that these varying demand conditions along the latitudinal dimension of the system trace out the supply curve for this particular technology.<sup>9</sup> Given fixed supply and geographical variation in demand, we can consistently estimate the parameters of the LCF for each lock technologies separately.

<sup>9</sup>The geographical variation is displayed in the dynamic Figure 6.



**Figure 5:** Scatterplot of lock costs against the number of barges locked, 1993–2010, indicating (1) a quadratic relationship between flow and congestion for the standard configuration; (2) double-chamber 600 ft locks, and (3) 1200 ft locks.



**Figure 6:** Dynamic scatterplot of lock costs against the number of barges locked, 1993–2010. Varying demand conditions along the latitudinal dimension of the system of locks trace out the supply curve or lock congestion function for 600 ft technology; this diagrams consecutively (upstream to downstream) plots all observations  $\{\text{barges}_{\ell ct}, \text{costs}_{\ell ct}\}_{t=1}^T$  for  $\ell \in \mathcal{L}$ .

## 4 Estimating Lock Congestion

### 4.1 Regression Specification

Given the above discussion on controlling for simultaneity bias through the exploitation of the individual locks' geographical location and intensity of use, and the description of lock characteristics and lock usage, we propose to estimate the following fixed-effects panel regression model separately for each type of lock technology:

$$\begin{aligned} \text{costs}_{\ell ct} = & \beta_1 \text{barges}_{\ell ct} + \beta_2 \text{barges}_{\ell ct}^2 + \boldsymbol{\zeta}' \mathbf{x}_{\ell ct} \\ & + \gamma_1 (\text{NCHMB}_{\ell ct} \times \text{barges}_{\ell ct}) \\ & + \gamma_2 (\text{NCHMB}_{\ell ct} \times \text{barges}_{\ell ct}^2) \\ & + \mu_{\ell c} + \tau_t + \varepsilon_{\ell ct}, \end{aligned} \quad (1)$$

for  $\ell \in \mathcal{L}_c$ ,  $c \in \mathcal{C} = \{600, 1200\}$  ft, and  $t \in \mathcal{T} = [1993, 2010]$ .

The dependent variable in equation (1),  $\text{costs}_{\ell ct}$ , is the average lock delay and processing time (in minutes) incurred at lock  $\ell \in \mathcal{L}_c$  in year  $t \in \mathcal{T}$ , where  $\mathcal{L}_c$  denotes the set of locks featuring 600 or 1200 ft technology. The main explanatory variable on the right-hand-side of equation (1),  $\text{barges}_{\ell ct}$ , is the number of barges locked (in thousands) at lock  $\ell \in \mathcal{L}_c$  in year  $t \in \mathcal{T}$ . To investigate the effect of the number of chambers on lock congestion, we include a binary variable  $\text{NCHMB}_{\ell ct}$  that equals one in the presence of an auxiliary chamber (and equals zero otherwise), and an interaction effect between the number of chambers and the (square of the) number of barges locked,  $\text{NCHMB}_{\ell ct} \times \text{barges}_{\ell ct}$ . We thus explicitly model the effect of the number of chambers as a slope and intercept effect.

Furthermore we include a vector of lock-specific control variables  $\mathbf{x}_{\ell ct}$  that might affect lock congestion; this vector includes age (in years), scheduled and unscheduled lock unavailabilities (measured by the number of lock outages in year  $t$ ), and lockages of non-commercial and recreational vessels (in thousands). Note that equation (1) accounts both for heterogeneity by including lock-specific fixed effects,  $\mu_{\ell c}$ , and aggregate shocks that affect all locks simultaneously through the inclusion of time fixed effects  $\tau_t$ .<sup>10</sup> Any remaining unobserved factors influencing lock congestion are captured by the lock-specific error term  $\varepsilon_{\ell ct}$ . The error term  $\varepsilon_{\ell ct}$  has zero conditional mean  $\mathbb{E}[\varepsilon_{\ell ct} | \mathbf{x}_{\ell ct}]$  and is uncorrelated with the explanatory variables,  $\text{Corr}\{\mathbf{x}_{\ell ct}, \varepsilon_{\ell ct}\} = 0$ .

### 4.2 Results

From equation (1) it follows that the parameter vector of interest is  $\boldsymbol{\theta} = (\beta_1, \beta_2, \gamma_1, \gamma_2, \boldsymbol{\zeta})'$ . While we have reason to believe  $\beta_1, \beta_2 > 0$ , subsequent regression estimates will reveal whether  $\gamma_1, \gamma_2 < 0$  is a reasonable hypothesis in the current context.

<sup>10</sup>A boxplot of lock delay costs (not shown here) by year indicates variation over time; the distribution of lock processing costs on the contrary appears stationary over time.

**Table 4:** OLS and FE estimates of 600 ft lock congestion function parameters

	I	II	III
(Intercept)	44.322 (105.217)		
barges <sub>lct</sub>	1.542 (1.330)	-1.394 (2.486)	-2.858 (4.363)
barges <sub>lct</sub> <sup>2</sup>	0.125*** (0.037)	0.132* (0.055)	0.151* (0.067)
NCHMB <sub>lct</sub>	93.793 (253.237)		
NCHMB <sub>lct</sub> × barges <sub>lct</sub>	-8.017 (23.987)	-9.316 (13.795)	-5.706 (18.418)
NCHMB <sub>lct</sub> × barges <sub>lct</sub> <sup>2</sup>	0.143 (0.541)	0.205 (0.341)	0.115 (0.450)
age	-0.012 (1.417)		
unavailUnsch	0.281* (0.121)	0.258‡ (0.172)	0.279† (0.163)
unavailSch	-0.358 (0.409)	0.091 (0.431)	-0.238 (0.367)
vesselsNonCommercial	32.980 (59.362)	74.965 (64.178)	53.323 (69.453)
vesselsRecreational	-2.459*** (0.578)	0.225 (1.498)	3.493‡ (2.475)
Time FE	No	No	Yes
R <sup>2</sup>	0.754	0.336	0.510
$\bar{R}^2$	0.747	0.310	0.450
N	396	396	396

HC4 robust standard errors in parentheses

Legend: \*\*\*  $p < 0.001$ , \*\*  $p < 0.01$ , \*  $p < 0.05$ , †  $p < 0.1$ , ‡  $p < 0.15$ 

Tables 4 and 5 summarize the estimates for our regression specification (1).<sup>11</sup> We discuss estimates for 600 and 1200 ft technology separately; for 600 ft technology columns I, II and III from Table 4 respectively report the OLS, lock FE, and lock and time FE estimates. Columns I and II from Table 5 report lock and time FE estimations. For the purposes of statistical inference, we rely on (and report) robust standard errors that control for heteroskedasticity of unknown form and leverage effects in the matrix of regressors (referred to as HC4 standard errors), see White (1980) and Cribari-Neto *et al.* (2007, 2008).

#### 4.2.1 600 ft Lock Technology

Columns I, II, and II from Table 4 report a statistically significant and positive effect of the square of the number of barges locked on lock congestion, providing evidence for the

<sup>11</sup>See Appendix A for more details on how the panel data set in this analysis was constructed.

existence of a [Bureau of Public Roads \(1964\)](#) style quadratic lock congestion function. Using short-hand notation  $c$  for costs and  $b$  for barges, the interpretation of the size of the coefficient on  $\hat{\beta}_2$  is that an increase in the number of barges by an amount  $\Delta b$  increases estimated lock congestion with an amount equal to  $\Delta \hat{c} = (\hat{\beta}_1 + 2\hat{\beta}_2 \bar{b})\Delta b$ , where  $\bar{b}$  can be evaluated at the lock mean or median. The marginal effect on lock congestion then equals

$$\frac{\partial \hat{c}}{\partial b} = \hat{\beta}_1 + 2\hat{\beta}_2 \bar{b} = 0 + 2 \begin{pmatrix} 0.125 \\ 0.132 \\ 0.152 \end{pmatrix} \bar{b}, \quad (2)$$

depending on the method of estimation. For example, with lock and time FE estimates (Column III) the impact of a marginal increase in the number of barges on congestion for a lock that on average handles 20,000 barges per year is  $2 \times 0.152 \times 20 = 6.08$  minutes.<sup>12</sup> Note that the estimated coefficient  $\hat{\beta}_2$  increases in magnitude when considering lock and time FE; while consistent (on the basis of our main identification strategy) the OLS estimate seems to underestimate the true effect of traffic on congestion when not controlling for unobserved heterogeneity.

Considering lock characteristics, there is no statistical evidence that having an additional chamber affects congestion for 600 ft technology, nor do the age of the lock, scheduled outages, or non-commercial vessels. The only control variables that affect congestion are unscheduled outages, and the number of recreational vessels handled throughout the year. Given the estimates in Table 4, 10 additional outages increase average congestion with 2.81, 2.58 and 2.79 minutes respectively.<sup>13</sup> Similarly, an additional 1000 recreational vessels results in an increase of 3.5 minutes of lock congestion time.

Under the FE framework, the estimated fixed effects  $\hat{\mu}_{\ell,c}$  have the interpretation of the average level of congestion for each lock when all covariates are equal to zero. Figure 7 shows the distribution of these lock fixed effects using a nonparametric kernel density estimation; from these graphs, we infer that (i)  $\hat{\mu}_{\ell,c} \in [35.19, 189.40]$  or  $\hat{\mu}_{\ell,c} \in [50.44, 256.84]$  minutes respectively, and (ii) the corresponding means are

$$\mathbb{E}[\hat{c}|b = 0, \mathbf{x}_{\ell ct} = \mathbf{0}] = \begin{pmatrix} 88.57 \\ 140.17 \end{pmatrix}$$

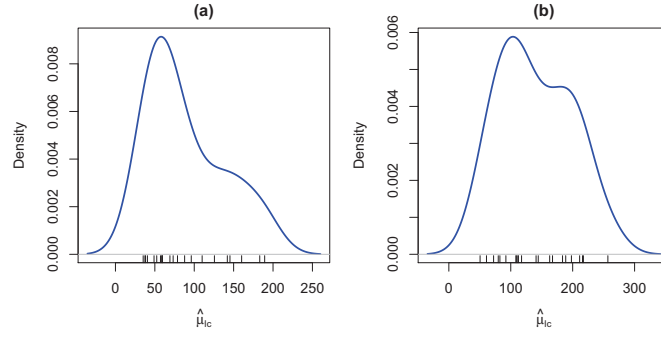
minutes under estimation methods II and III.

#### 4.2.2 1200 ft Lock Technology

Opposed to the findings for 600 ft technology, there is no statistical evidence of a *direct* effect of barges on lock congestion as we fail to reject the null hypothesis  $\hat{\beta}_1, \hat{\beta}_2 = 0$ , see

<sup>12</sup>Since  $\text{barges}_{\ell ct}$  is measured in thousands of units, the marginal increase corresponds here with an additional 1,000 barges of traffic.

<sup>13</sup>The fact that unscheduled lock outages become marginally significant when using lock and time FE might be purely due to limited sample size; additional observations would lead to sharper estimation of the corresponding coefficients.



**Figure 7:** Distribution of estimated lock fixed effects  $\hat{\mu}_{\ell,c}$  from equation (1) for  $c = 600$  ft lock technology; nonparametric kernel density estimation. Panel (a) without time fixed effects; panel (b) including time fixed effects.

Table 5.<sup>14</sup> This finding suggest that, given the data and the model specified, 1200 ft locks (with and without auxiliary chambers) operate under free-flow conditions. The important result from the estimates in Table 5 is the statistically significant effect of the number of lock chambers on lock congestion. The impact of a marginal increase in traffic is (in vector notation)

$$\begin{aligned} \frac{\partial \hat{c}}{\partial \bar{b}} &= \hat{\beta}_1 + 2\hat{\beta}_2\bar{b} + \hat{\gamma}_1\text{NCHMB}_{\ell ct} + 2\hat{\gamma}_2(\text{NCHMB}_{\ell ct} \times \bar{b}) \\ &= 0 + 0 + \begin{pmatrix} 30.365 \\ 55.465 \end{pmatrix} + 2 \begin{pmatrix} -0.307 \\ -0.521 \end{pmatrix} \bar{b}, \end{aligned} \quad (3)$$

where the second equality follows from the estimates in Table 5. Result (3) implies that an increase in the number of barges for double-chamber 1200 ft locks leads to more lock costs, but these costs are decreasing with the level of demand. For example, given the size of traffic at these particular locks, with an average at 65.000 barges a year, these interaction effects amount to  $\frac{\partial \hat{c}}{\partial \bar{b}} \in \{-9.545, -12.265\}$  minutes respectively, depending on the estimation method I or II.<sup>15</sup>

We further see that lock outages significantly affect lock congestion times: an additional unscheduled lock outage adds and extra 1.044 or 1.244 minutes to the average lock congestion time. Surprisingly, non-commercial vessel traffic seems to have a tempering effect on lock congestion.<sup>16</sup>

<sup>14</sup>Singularity issues prevented us from retrieving the fixed effects for all but one lock.

<sup>15</sup>This means that for an additional 1000 barges handled, lock congestion costs decrease with  $\{-9.545, -12.265\}$  minutes for a double-chamber 1200 ft lock with respect to a single-chamber lock of the same length. The effect is calculated as  $30.365 - 2 \times 0.307 \times 65 = -9.545$  and  $55.465 - 2 \times 0.521 \times 65 = -12.265$ , as the explanatory variable is expressed in thousands of units.

<sup>16</sup>If this type of traffic includes dredging ships or maintenance vessels for example, this result might not be that surprising at all.

**Table 5:** FE estimates of 1200 ft lock congestion function parameters

	I	II
$\text{barges}_{\ell_{ct}}$	-7.579 (7.923)	-28.952 (22.705)
$\text{barges}_{\ell_{ct}}^2$	0.147 (0.153)	0.272 (0.448)
$\text{NCHMB}_{\ell_{ct}} \times \text{barges}_{\ell_{ct}}$	30.365* (13.155)	55.465*** (10.962)
$\text{NCHMB}_{\ell_{ct}} \times \text{barges}_{\ell_{ct}}^2$	-0.307* (0.146)	-0.521* (0.211)
$\text{unavail}_{\text{Unsch}}$	1.044** (0.414)	1.244** (0.390)
$\text{unavail}_{\text{Sch}}$	1.482 (7.132)	0.283 (2.173)
$\text{vessels}_{\text{NonCommercial}}$	-409.465 (495.861)	-291.694* (114.177)
$\text{vessels}_{\text{Recreational}}$	-0.454 (4.183)	-14.439 (20.819)
Time FE	No	Yes
$R^2$	0.202	0.461
$\bar{R}^2$	0.161	0.222
N	54	54

HC4 robust standard errors in parentheses

Legend: \*\*\*  $p < 0.001$ , \*\*  $p < 0.01$ , \*  $p < 0.05$ , †  $p < 0.1$ , ‡  $p < 0.15$ 

## 5 Conclusion

In this paper we argued that the geographical variation in demand patterns together combined with a fixed supply technology serves as an instrument to consistently estimate the parameters of a congestion function for locks contained within a river system of locks. Applying the argument to a compiled data set of traffic and congestion in locks on the Upper Mississippi River system of locks and dams we find compelling statistical evidence of the existence of a quadratic relationship between river barge traffic and lock congestion for 600 ft lock technology, controlling for lock characteristics and unobserved heterogeneity. While we find no such direct effect for 1200 ft lock technology, implying that these locks operate under free-flow conditions, the presence of an auxiliary lock chamber significantly mitigates lock congestion for this type of technology, and this effect is more pronounced for higher levels of traffic.

Irrespective of the technology used, unscheduled lock outages negatively affect lock congestion times; it is however a relief to see that shippers behave rationally and take scheduled outages into account, as corroborated by a statistically insignificant coefficient on the corresponding covariate. Furthermore, an increase in the amount of traffic generated by recreational vessels negatively affects lock congestion on 660 ft technology, while surprisingly, more non-commercial vessels decrease lock costs on 1200 ft locks.

The empirical findings in this paper corroborate the rather subjective assessment that 600 ft lock technology is no longer adequate to deal with current-day levels of barge traffic on the Upper Mississippi, and provide evidence in favor of replacement with 1200 ft, double-chamber lock technology so as to mitigate lock congestion. The present methodology can be applied to provide input under the form of parameters of the LCF in traffic assignment problems, and in studies in need of consistent estimates of demand and supply for inland waterway transport, or given these, to assess the welfare effects accruing from improvements to lock infrastructure.

## References

- Boland, J.J., Brezonik, P., Davis, R.K., Eisel, L.L., Fuller, S.W., Galloway, G.E., Lave, L.B., Limburg, K., Rieke, E.E., Sorooshian, S. and Sparks, R.E. (2004), “Review of the U.S. Army Corps of Engineers Restructured Upper Mississippi River-Illinois Waterway Feasibility Study—Second Report,” *Tech. rep.*, National Research Council, National Academy of Sciences, Washington, DC. [1, 3]
- Bureau of Public Roads (1964), *Traffic Assignment Manual*, Urban Planning Division, U.S. Department of Commerce, Washington, DC. [11]
- Cribari-Neto, F., Souza, T. and Vasconcellos, K. (2007), “Inference under Heteroskedasticity and Leveraged Data,” *Communications in Statistics – Theory and Methods* 36(10), 1877–1888. [10]
- (2008), “Inference under Heteroskedasticity and Leveraged Data. Errata,” *Communications in Statistics – Theory and Methods* 37(20), 3329–3330. [10]
- Croissant, Y. and Millo, G. (2008), “Panel Data Econometrics in R: The plm Package,” *Journal of Statistical Software* 27(2), URL <http://www.jstatsoft.org/v27/i02/>. [17]
- Deweese, D.N. (1979), “Estimating the Time Costs of Highway Congestion,” *Econometrica* 47(6), 1499–1512. [3]
- Hayashi, F. (2000), *Econometrics*, Princeton University Press, Princeton, New Jersey. [2]
- Hazelton, M.L. and Pueschel, J. (1999), “Estimation of Link Performance Functions from Incomplete Flow Data,” *Journal of Advanced Transportation* 33(3), 323–334. [3]
- Johnson, M.B. (1964), “On the Economics of Road Congestion,” *Econometrica* 32(1-2), 137–150. [3]
- Leamer, E.E. (1981), “Is It a Demand Curve, or Is It a Supply Curve? Partial Identification through Inequality Constraints,” *The Review of Economics and Statistics* 63(3), 319–327. [1]



- R Core Team (2012), *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, URL <http://www.R-project.org/>, ISBN 3-900051-07-0. [17]
- Rose, G. (1994), “Impacts of Estimation Errors in Travel Time Functions,” *Transportation Planning and Technology* 18(4), 249–262. [3]
- Rose, G. and Raymond, M. (1992), “Simulated Estimation of Davidson’s Travel Time Function,” *Transportation Planning and Technology* 16(4), 251–259. [3]
- Russo, F. and Vitetta, A. (2011), “Reverse Assignment: Calibrating Link Cost Functions and Updating Demand from Traffic Counts and Time Measurements,” *Inverse Problems in Science and Engineering* 19(7), 921–950. [3]
- Sheffi, Y. (1985), *Urban Transportation Networks. Equilibrium Analysis with Mathematical Programming Methods*, Prentice-Hall, Englewood Cliffs, New Jersey. [2]
- Taylor, M. (1977), “Parameter Estimation and Sensitivity of Parameter Values in a Flow-Rate/Travel-Time Relation,” *Transportation Science* 11(43), 275–292. [3]
- USACE (1993), “Compendium on Waterway Transportation Reliability: Lock Congestion and Lock Queues,” *Iwr report 93-R-9*, US Army Corps of Engineers, Water Resources Support Center, Institute for Water Resources, Alexandria, VA, URL [www.iwr.usace.army.mil/docs/iwrreports/93-R-9.pdf](http://www.iwr.usace.army.mil/docs/iwrreports/93-R-9.pdf). [3]
- (2002), “Interim Report for the Restructured Upper Mississippi River-Illinois Waterway System Navigation Feasibility Study,” *Report*, US Army Corps of Engineers, Alexandria, VA, <http://www2.mvr.usace.army.mil/umr-iwwsns/>. [1]
- (2004), “Integrated Feasibility Report and Programmatic Environmental Impact Statement for the UMR-IWW System Navigation Feasibility Study,” *Report*, US Army Corps of Engineers, Alexandria, VA, <http://www2.mvr.usace.army.mil/umr-iwwsns/>. [1]
- (2011a), “Lock Performance Monitoring System (LPMS),” *Data repository*, US Army Corps of Engineers, Institute for Water Resources, Alexandria, VA, URL <http://www.ndc.iwr.usace.army.mil//lpms/lpms.htm>. [5, 17, 20]
- (2011b), “Lock Performance Monitoring System (LPMS): Definition of Terms,” *Data repository*, US Army Corps of Engineers, Institute for Water Resources, Alexandria, VA, URL [http://www.ndc.iwr.usace.army.mil/lpms/pdf/lpmsstat\\_v3.pdf](http://www.ndc.iwr.usace.army.mil/lpms/pdf/lpmsstat_v3.pdf). [17]
- (2011c), “Navigation Data Center (NDC),” *Data repository*, US Army Corps of Engineers, Institute for Water Resources, Alexandria, VA, URL <http://www.ndc.iwr.usace.army.mil/index.htm>. [17, 20]

- (2011*d*), “U.S. Waterway Data: Lock Characteristics,” *Data repository*, US Army Corps of Engineers, Institute for Water Resources, Alexandria, VA, URL <http://www.ndc.iwr.usace.army.mil/data/datalck.htm>. [4, 17, 20]
- (2012), “Upper Mississippi River-Illinois Waterway System Navigation Study Website,” <http://www2.mvr.usace.army.mil/umr-iwwsns/>, consulted 20/08/2012. [1]
- Walters, A.A. (1961), “The Theory and Measurement of Private and Social Cost of Highway Congestion,” *Econometrica* 29(4), 676–699. [3]
- White, H. (1980), “A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity,” *Econometrica* 48(4), 817–838. [10]
- Working, E. (1927), “What Do Statistical “Demand Curves” Show?” *The Quarterly Journal of Economics* 41(2), 212–235. [1]
- Working, H. (1925), “The Statistical Determination of Demand Curves,” *The Quarterly Journal of Economics* 39(4), 503–543. [1]
- Zeileis, A. (2004), “Econometric Computing with HC and HAC Covariance Matrix Estimators,” *Journal of Statistical Software* 11(10), 1–17, URL <http://www.jstatsoft.org/v11/i10/>. [17]
- (2006), “Object-oriented Computation of Sandwich Estimators,” *Journal of Statistical Software* 16(9), 1–16, URL <http://www.jstatsoft.org/v16/i09/>. [17]

## A Data Collection and Computational Details

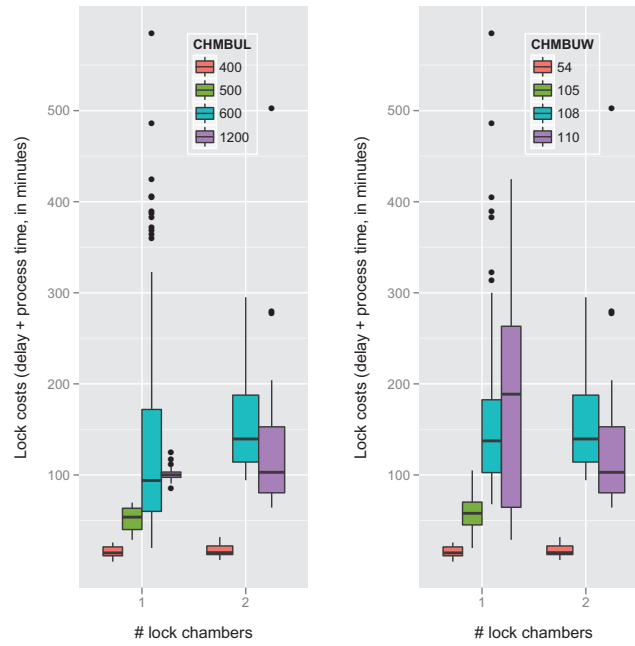
All data on lock usage were obtained from the publicly available section of the Navigation Data Center maintained by the U.S. Army Corps of Engineers, see [USACE \(2011c\)](#). Specifically, data on lock usage, lock outages and commodities locked (tonnages) were obtained from the *Lock Performance Monitoring System* (LPMS) ([USACE, 2011a](#)), while numbers for lock characteristics were extracted from the U.S. Waterway Data section ([USACE, 2011d](#)). For a definition of the aggregates used, see [USACE \(2011b\)](#).

Data preparation and analysis for this paper was performed using the R language and environment for statistical computing and graphics, specifically R version 2.15.1 “Roasted Marshmallows” ([R Core Team, 2012](#)). Computing backend: Dell OptiPlex 980 equipped with an Intel Core i7 CPU with 16 GB RAM. Operating system: 64-bit Linux Ubuntu version 11.04 “Natty.” Fixed-effects regressions were performed using the **plm** package ([Croissant and Millo, 2008](#)), and HC4 robust standard errors were computed with the **sandwich** package ([Zeileis, 2004, 2006](#)).

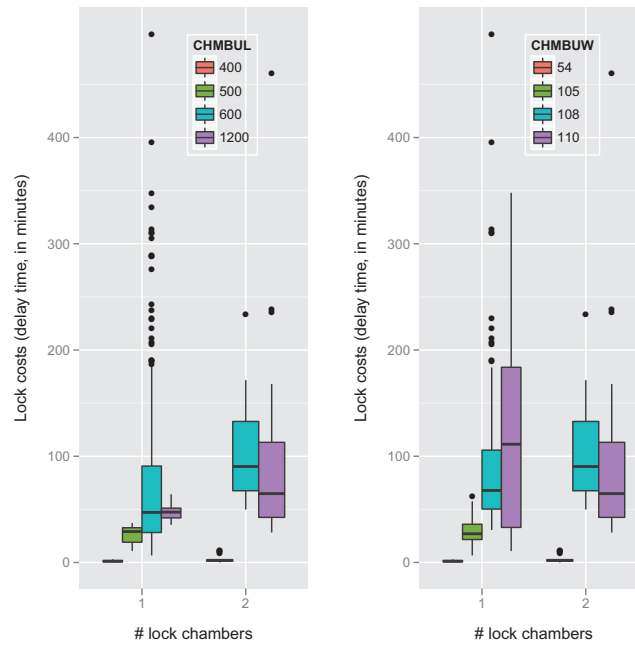
## B Supplementary Tables and Figures

**Table 6:** USACE commodity group classification

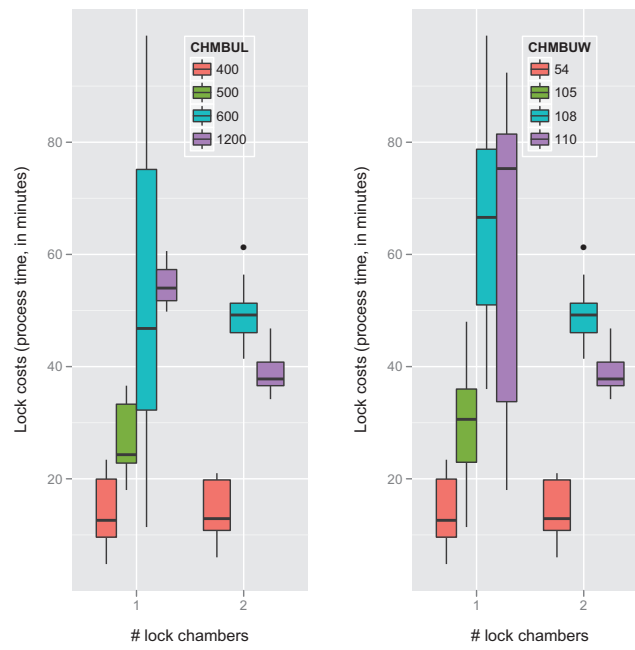
Code	Commodity group
10	coal, lignite, and coal coke
20	petroleum and petroleum products
30	chemicals and related products
40	crude materials, inedible, except fuels
50	primary manufactured goods
60	food and farm products
70	manufactured equipment & machinery
80	waste material
90	unknown or not elsewhere classified



**Figure 8:** Boxplots of lock costs (delay + proces time) by number of chambers and chamber useable length (CHMBUL, left panel) and chamber useable width (CHMBUW, right panel).



**Figure 9:** Boxplots of lock delay time by number of chambers and chamber useable length (CHMBUL, left panel) and chamber useable width (CHMBUW, right panel).



**Figure 10:** Boxplots of lock processing time by number of chambers and chamber useable length (CHMBUL, left panel) and chamber useable width (CHMBUW, right panel).

**Table 7: Upper Mississippi River Lock Characteristics<sup>a</sup>**

lock	id	code <sup>b</sup>	# <sup>c</sup>	chamber	mile	year	length	length <sup>d</sup>	width	width <sup>e</sup>	lift <sup>f</sup>	status	gate	latitude	longitude	state	town	county
27	29	3017938601	2	main	185.50	1953	1200	1200	110	110	21	ops	vertical	38.70	-90.18	IL	Granite City	Madison
27	29	3017938604	2	aux	185.50	1953	600	600	110	110	21	ops	miter	38.70	-90.18	IL	Granite City	Madison
mp	28	3020090601	2	main	200.80	1990	1200	1200	110	110	24	ops	vertical	38.87	-90.15	IL	East Alton	Madison
mp	28	3020090604	2	aux	200.80	1994	600	600	110	110	24	ops	miter	38.87	-90.15	IL	East Alton	Madison
25	27	3024140601	1	main	241.40	1939	600	600	110	110	15	ops	miter	39	-90.69	MO	Winfield	Lincoln
24	26	3027340601	1	main	273.40	1940	600	600	110	110	15	ops	miter	39.38	-90.91	MO	Clarksville	Pike
22	25	3130120601	1	main	301.20	1938	600	600	110	108	10	ops	miter	39.64	-91.25	MO	Saverton	Ralls
21	24	3132410001	1	main	324.90	1938	600	600	110	108	10	ops	miter	39.91	-91.43	IL	Quincy	Adams
20	23	3134310001	1	main	343.20	1936	600	600	110	108	10	ops	miter	40.14	-91.51	MO	Canton	Lewis
19	22	3136410001	1	main	364.30	1957	1200	1200	110	108	38	ops	vertical	40.39	-91.37	IA	Keokuk	Lee
18	21	3141050601	1	main	410.50	1937	600	600	110	108	10	ops	miter	40.88	-91.02	IL	Gladstone	Henderson
17	20	3143710001	1	main	437.10	1939	600	600	110	108	8	ops	miter	41.19	-91.06	IL	New Boston	Mercer
16	19	3145710001	1	main	457.20	1937	600	600	110	108	9	ops	miter	41.43	-91.01	IL	Muscatine	Rock Island
15	18	3148210001	2	main	482.90	1934	600	600	110	108	16	ops	miter	41.52	-90.57	IA	Rock Island	Rock Island
15	18	3148210004	2	aux	482.90	1934	360	360	110	108	16	ops	miter	41.52	-90.57	IL	Rock Island	Rock Island
14	17	3149310601	2	main	493	1922	600	600	110	108	11	ops	miter	41.57	-90.40	IA	LeClaire	Scott
14	17	3149310604	2	aux	493	1939	320	320	80	78	11	season	miter	41.57	-90.40	IA	LeClaire	Scott
13	16	3152210001	1	main	522.50	1938	600	600	110	108	11	ops	miter	41.90	-90.16	IL	Clinton	Whiteside
12	15	3155610001	1	main	556.70	1939	600	600	110	108	9	ops	miter	42.26	-90.42	IA	Bellevue	Jackson
11	14	3158310001	1	main	583	1937	600	600	110	108	12	ops	miter	42.54	-90.64	IA	Dubuque	Dubuque
10	13	3261510601	1	main	615.10	1936	600	600	110	105	8	season	miter	42.78	-91.09	IA	Guttenburg	Clayton
9	12	3264790601	1	main	647.90	1938	600	600	110	105	9	season	miter	43.21	-91.10	WI	Lynxville	Crawford
8	11	3267920601	1	main	679.20	1937	600	600	110	105	11	season	miter	43.57	-91.23	WI	Genoa	Vernon
7	10	3270250601	1	main	702.50	1937	600	600	110	105	8	season	miter	43.87	-91.31	MN	Dresbach	Winona
6	9	3271430601	1	main	714.30	1936	600	600	110	105	6	season	miter	44	-91.44	WI	Trempealeau	Trempealeau
5a	8	3272850601	1	main	728.50	1936	600	600	110	105	5	season	miter	44.09	-91.67	MN	Winona	Winona
5	7	3273810601	1	main	738.10	1935	600	600	110	105	9	season	miter	44.16	-91.81	MN	Minneiska	Winona
4	6	3275280601	1	main	752.80	1935	600	600	110	105	7	season	miter	44.33	-91.92	WI	Alma	Buffalo
3	5	3279690601	1	main	796.90	1938	600	600	110	105	8	season	miter	44.61	-92.61	MN	Red Wing	Goodhue
2	4	3281520601	1	main	815.20	1930	500	500	110	110	12	season	miter	44.76	-92.87	MN	Hastings	Dakota
1	3	3284060601	1	main	847.60	1930	400	400	56	54	38	season	miter	44.92	-93.20	MN	St. Paul	Hennepin
lsaf	2	4001940602	1	main	853.30	1959	400	400	56	54	25	season	miter	44.98	-93.25	MN	Minneapolis	Hennepin
usaf	1	3285310101	1	main	853.90	1963	400	400	56	54	49	season	miter	44.98	-93.26	MN	Minneapolis	Hennepin

<sup>a</sup> Source: adapted from [USACE \(2011a,c,d\)](#)

<sup>b</sup> NDCCODE: unique Navigation Data Center (NDC) dock & lock ID

<sup>c</sup> NOCHMB: number of chambers in the lock

<sup>d</sup> CHMBUL: chamber useable length (in ft)

<sup>e</sup> CHMBUW: chamber useable width (in ft)

<sup>f</sup> LIFT: distance from empty to full (in ft)